Published in final edited form as:
Structure. 2006 January ; 14(1): 63–73. doi:10.1016/j.str.2005.07.025.

Mapping the Structure and Function of the E1 and E2 Glycoproteins in Alphaviruses

Suchetana Mukhopadhyay[1,3,*], **Wei Zhang**[1], **Stefan Gabler**[1], **Paul R. Chipman**[1], **Ellen G. Strauss**[2], **James H. Strauss**[2], **Timothy S. Baker**[1,4], **Richard J. Kuhn**[1], and **Michael G. Rossmann**[1]

[1] Department of Biological Sciences, Purdue University, 915 West State Street, West Lafayette, Indiana 47907

[2] Division of Biology, California Institute of Technology, Mail Code 156-29, 1200 East California Boulevard, Pasadena, California 91125

Summary

The 9 Å resolution cryo-electron microscopy map of Sindbis virus presented here provides structural information on the polypeptide topology of the E2 protein, on the interactions between the E1 and E2 glycoproteins in the formation of a heterodimer, on the difference in conformation of the two types of trimeric spikes, on the interaction between the transmembrane helices of the E1 and E2 proteins, and on the conformational changes that occur when fusing with a host cell. The positions of various markers on the E2 protein established the approximate topology of the E2 structure. The largest conformational differences between the icosahedral surface spikes at icosahedral 3-fold and quasi-3-fold positions are associated with the monomers closest to the 5-fold axes. The long E2 monomers, containing the cell receptor recognition motif at their extremities, are shown to rotate by about 180° and to move away from the center of the spikes during fusion.

Introduction

Cryo-electron microscopy (cryo-EM) reconstructions of various alphaviruses have been reported with progressively improved resolution (Cheng et al., 1995; Mancini et al., 2000; Paredes et al., 1993, 2001; Zhang et al., 2002a, 2002b). Alphavirus particles are icosahedral structures with an external diameter of about 700 Å, and they contain a lipid bilayer derived from the host plasma membrane and one copy of an ~12 kb positive-strand, genomic RNA. Five structural proteins—capsid, E3, E2, 6K, and E1—are translated during virus replication. Each alphavirus particle contains 240 copies each of the E1 (439 amino acid residues; Sindbis virus numbering), E2 (423 amino acid residues), and capsid proteins (264 amino acid residues), all of which are arranged with T = 4 quasi-symmetry (Caspar and Klug, 1962). E1 and pE2 (precursor to the E3 and E2 proteins before cleavage of E3) are assembled as heterodimers in the endoplasmic reticulum, E3 is cleaved in the Golgi, and the resultant E1-E2 heterodimers are then transported to the plasma membrane. These heterodimers, to which E3 may or may not remain associated, depending on the virus, self-assemble into 80 trimeric spikes on the

* Correspondence: sumukhop@indiana.edu.
[3]Present address: Department of Biology, Indiana University, 1001 East Third Street, Bloomington, Indiana 47405.
[4]Present address: Department of Chemistry and Biochemistry, University of California, San Diego, 9500 Gilman Drive #0378, La Jolla, California 92093.

Accession Numbers: The 9 Å cryo-EM map has been deposited with the European Bioinformatics Institute under accession code EMD-1121. The coordinates of the fitted ectodomain E1, transmembranes E1 and E2, and the capsid protein have been deposited with the Protein Data Bank under accession code 1Z8Y.

virus surface (von Bonsdorff and Harrison, 1975, 1978). E1 is responsible for cell fusion, and E2 is primarily involved in receptor binding and cell entry. The function of E3 is unclear. E1 and E2 each have one transmembrane helix that traverses the lipid bilayer. The lipid bilayer surrounds the nucleocapsid core, which is formed by the capsid protein together with the genomic RNA. The small 6K protein (6 kDa) associates with the E1-pE2 heterodimer and is transported to the plasma membrane with the E1-E2 heterodimer. The stoichiometry of 6K to E1-pE2 is 1:1, but in virions the proportion of 6K is considerably lower (Gaedigk-Nitschko and Schlesinger, 1990; Lusa et al., 1991). The presence and location of 6K have yet to be identified in any of the cryo-EM virion structures. However, there are thought to be two successive hydrophobic regions in the 6K protein that form antiparallel helices within the membrane.

The X-ray crystal structures of the ectodomain of the E1 protein (residues 1–383) of Semliki Forest virus (SFV) and the carboxy-terminal region of the capsid protein (residues 114–264) of Sindbis virus (SINV) have been determined (Choi et al., 1991; Lee et al., 1996; Lescar et al., 2001). The E1 ectodomain consists of three β barrel domains (Lescar et al., 2001). Domain I contains the amino terminus and is spatially located between domains II and III. The carboxy terminus lies within domain III, and the fusion peptide is at the distal end of domain II. The amino-terminal region of the capsid protein (residues 1–113) is apparently disordered in crystals and therefore not visible in the X-ray structure (Choi et al., 1991). However, the carboxy-terminal region is folded into a chymotrypsin-like fold, and the protein acts as a proteinase during virus assembly. The carboxy-terminal domain contains a hydrophobic pocket that binds the short (33 amino acids) cytoplasmic region of E2 (Lee et al., 1996; Skoging et al., 1996).

The atomic structures of the E1 ectodomain and the capsid protein have been fitted into an 11 Å cryo-EM reconstruction of SINV, generating a partial pseudoatomic structure of the virus (Zhang et al., 2002b). The E1 monomers were found to lie at the base of each of the surface spikes and to form a trimer around each of the icosahedral and quasi-3-fold axes, resulting in the formation of a lattice on the virus surface. Domain III of E1 lies closest to the lipid bilayer, and domains I and II, which protrude away from the lipid bilayer, constitute the underside of the spike. One heterodimer from the 3-fold spike and three heterodimers from a quasi-3-fold spike are related to each other by T = 4 quasi-symmetry (Figure 1). After the densities corresponding to the E1 molecules (as determined from the pseudoatomic structure) (Zhang et al., 2002b) had been subtracted from the cryo-EM map, the remaining density outside of the phospholipid bilayer could be attributed mainly to the E2 glycoprotein. This density was found to be a long, thin molecule that covers the top of the trimeric E1 molecules.

The atomic structure of E2 has yet to be determined. However, a number of sites along the polypeptide have been identified and thus provide some structural information (Paredes et al., 1998; Pletnev et al., 2001; Smith et al., 1995; Zhang et al., 2005). Escape mutants to neutralizing monoclonal antibodies against E2 (Davis et al., 1987; Meyer and Johnston, 1993; Strauss et al., 1991) have suggested that amino acids 180–220 (SINV) are probably located on the virus surface. The cryo-EM reconstruction of Ross River virus (RRV) complexed with Fab fragments of one of these neutralizing antibodies led to an approximate location of residue E2-216, the site of an escape mutation (Smith et al., 1995). Similarly, residue E2-218, a residue essential for heparin binding, was located in a cryo-EM reconstruction of a mutant form of RRV complexed with heparin (Heil et al., 2001; Zhang et al., 2005). In addition, cryo-EM reconstructions of RRV, and several glycosylation site mutants of SINV, allowed the carbohydrate moieties associated with residues E2-196, E2-200, E2-262, and E2-318 to be identified (Pletnev et al., 2001). The approximate location of the N terminus of E2 was established by computing a difference map between SFV (which contains E3) and SINV (which lacks E3), consistent with the results of Paredes et al. (1998).

Here, we report a 9 Å resolution cryo-EM reconstruction of SINV. The individual transmembrane densities representing the E1 and E2 glycoproteins could be fitted with atomic structures of α helices, thus permitting the E2 polypeptide to be traced from the surface spike, through the phospholipid bilayer, to its binding site on the capsid protein. Several additional glycosylated residues in the E2 ectodomain have been located, and these, together with the previously identified markers, determine the polypeptide topology of the entire E2 structure. Furthermore, our improved knowledge of the E2 structure provides information about the contacts between E1 and E2 within a heterodimer and shows that these contact regions on E1 become exposed on the surface of the postfusion complex (Gibbons et al., 2004).

Results and Discussion

Ectodomain of E1

A SINV E1 homology model based on the SFV E1 ectodomain X-ray structure was fitted into the 9 Å resolution cryo-EM density (Rossmann, 2000). Initially, the entire SINV E1 model was fitted as a rigid body by using the previously determined T = 4 quasi-symmetry operators (Table 1) (Zhang et al., 2002b). However, the quality of the fit improved when the four E1 monomers were fitted independently. The fit was further improved when each E1 molecule was treated as two rigid bodies, the first encompassing domains I+II (residues 1–290) and the second being domain III (residues 295–383) (Table 2). Domains I+II were fitted by using previously determined glycosylation sites for E1-N139, N141, and N245 as restraints (Pletnev et al., 2001). Domain III was then fitted into the density by restraining the position of its N terminus (residue 295) to be within 20 Å of the C terminus of domain I (residue 290). This independent fitting of two rigid bodies per E1 molecule led to a pseudoatomic model in which the four E1 monomers adopt slightly different conformations when comparing the spikes at the icosahedral 3-fold and quasi-3-fold axes (Figure 2). Notwithstanding the increased number of fitting parameters used in this procedure, there are small, yet significant, differences between the conformations of the E1 molecules at the icosahedral and the quasi-3-fold spikes.

Lescar et al. (2001) interpreted a homologous SFV map (Mancini et al., 2000) by using similar techniques, but they divided the E1 molecule into domains I+III as one rigid body and domain II as the other rigid body, as opposed to domains I+II and domain III as reported here. The Lescar et al. division of the E1 molecule produced a distinct worsening of the fit into the present 9 Å resolution map of SINV as measured by a decrease in the *sumf* values for all three domains in all four of the quasi-related positions.

Superposition of domains I+II of the four quasi-symmetric E1 proteins shows the variation in the position of the domains of each E1 monomer. Domain III of the monomer closest to the 5-fold axis (monomer 6 in Figure 1) has an rms deviation of 3.35, 2.49, and 3.21 Å with respect to monomers 1, 4, and 5. In contrast, the average rms displacement of domain III among monomers 1, 4, and 5 is 1.5 Å. The significance of the difference between the monomer 6 and monomer 1 structures is shown by the drop in the value of *sumf* when fitting monomer 6 into the position occupied by monomer 1. Presumably, the variability in the E1 monomer structure reflects differences in the packing constraints for the five monomers in close contact around each icosahedral 5-fold axis versus the six monomers around each icosahedral 2-fold axis (quasi-6-fold axis) (Figure 2). Furthermore, most of the E1-E1 contacts in the virion are between domain I of one monomer and domain III of the adjacent monomer at the icosahedral 5-fold axes. In contrast, contacts between the other E1 monomers are rather few.

In flaviviruses, the E fusion protein, which is structurally and functionally homologous to the E1 fusion protein of alphaviruses, undergoes significant conformational changes during both the maturation and fusion processes. The E protein is organized as trimers in immature flavivirus particles. Maturation of these particles into infectious virions requires large

conformational changes in the E protein that results in the loss of trimers and the formation of dimers. These changes involve an ~30° hinge motion between domains I and II (Zhang et al., 2004). Furthermore, as an initial step during the fusion process, the E protein has been postulated to undergo a rearrangement from homodimers to homotrimers, forming an intermediate fusion transition structure (Kuhn et al., 2002) (Figure 3). Formation of the postfusion trimer in both alpha- and flaviviruses requires a large movement at the hinge between domain I and domain III (Bressanelli et al., 2004; Gibbons et al., 2004; Modis et al., 2004). Notably, the structure of mature SINV is comparable to the fusion transition structure of flaviviruses. This could explain why the primary hinge movement observed in flaviviruses between immature and mature virus particles is between domains I and II, whereas the hinge movement observed in alphaviruses (corresponding to a fusion transition state in flaviviruses) is between domains I and III.

Ectodomain of E2

Density corresponding to the E2 glycoprotein and the E1 stem region in the 9 Å cryo-EM map was identified by setting to zero the density corresponding to the fitted ectodomain of the E1 molecules. The E2 density had a long, thin shape with a leaf-like structure at the top of the spike. There are a large number of contacts between the leaf-like structure of E2 and the polar residues at the distal end of the E1 glycoprotein domain II, where the fusion peptide is located, as well as between the stalk portion of E2 at the base of the spike and the hydrophobic residues in domains I and III of E1.

SINV mutants with extra glycosylation sites were generated either by site-directed mutagenesis (E2-Q46N, E2-E160N) or by multiple passages of virus in cell culture (E2-E216N [Strauss et al., 1991]). Cryo-EM difference maps between wild-type and mutant viruses identified the location of the additional carbohydrate moieties. These were validated by their T = 4 symmetry (Figure 4). The height of the difference density peaks (Pletnev et al., 2001; Zhang et al., 2002a, 2005) were 7.6σ, 8.4σ, and 11.2σ for E2-46, E2-160, and E2-216, respectively, where σ is the rms deviation of the density from the mean in the corresponding difference map. The carbohydrate moiety associated with residue E2-46 was located on the 3-fold axis of each spike at a viral radius of 310 Å, close to a neighboring E2 molecule. The difference density was smeared along each spike's 3-fold axis, possibly due to steric hindrance between adjacent carbohydrate groups. The carbohydrate associated with residue E2-160 mapped to the external side of the E2 leaf at a radius of 330 Å. The carbohydrate associated with E2-216 was found at the tip of the distal end of the spike on the leaf portion of E2, close to E2-218. This location is also the binding site for heparan sulfate in an RRV mutant (Zhang et al., 2005), as well as the Fab binding site for SINV and RRV neutralizing antibodies (Smith et al., 1995), consistent with the prediction that the residues E2-180 and E2-220 are surface accessible (Davis et al., 1987; Meyer and Johnston, 1993; Strauss et al., 1991). Recently, it has been suggested that residue E2-250 is at the E1-E2 interface (Liao and Kielian, 2005), which would also be in agreement with the current model of the E2 structure. The above-described results, together with the data reported earlier (Pletnev et al., 2001), show that the first 260 amino acids of E2 constitute the ectodomain, and are followed by about 100 amino acids that form the stem region before entering the lipid bilayer, crossed by a 30 amino acid-long helix. Finally, the carboxy-terminal domain of E2 consists of about 33 amino acids that interact with the nucleocapsid core.

It was previously suggested, based on secondary structure predictions and the shape of the E2 density, that E1 and E2 might have homologous folds, albeit with a rearrangement of domains along the polypeptide (Zhang et al., 2002b). This would place domain III of E2 into the leaf region at the top of the spike. As indicated above, this region is implicated in receptor binding and would be equivalent to domain III in the homologous flavivirus E protein, which also has

a probable receptor binding function (Crill and Roehrig, 2001; Mandl et al., 2000; Rey et al., 1995).

Movement of E1 and E2 during Fusion

Fusion in alphaviruses requires the dissociation of the E1-E2 heterodimer and formation of E1 homotrimers. The organization of the E1 trimers at the 3-fold and quasi-3-fold axes suggests that the three E1 molecules within a spike undergo conformational changes to form the homotrimers during fusion, ultimately forming the postfusion trimer (Gibbons et al., 2003, 2004). Residues of E1 that were in contact with E2 in the prefusion trimer end up located on the surface of the postfusion E1 trimer, facing away from the spike axes, and are predominantly in surface loops (Figure 5). Similarly, residues on the outside of the prefusion E1 trimer end up contacting other E1 molecules in the postfusion trimer. Thus, during the fusion process, the E1-E2 heterodimers rotate about their long axes, thereby moving the E2 molecules out of the center of the prefusion trimer and allowing the postfusion trimer to form.

The postfusion structures of the flavivirus E and alphavirus E1 glycoprotein trimers are similar (Bressanelli et al., 2004; Gibbons et al., 2004; Modis et al., 2004). Residues of the flavivirus E molecule that make E-E contacts in the postfusion structure are located on the internal surface of the E protein trimers of the fusion transition structure. Thus, in contrast to alphaviruses, there is no requirement for the rotation of the E glycoprotein about its long axis during the fusion process in flaviviruses. In alphaviruses, the functions of receptor recognition and fusion are attributed to E2 and E1, respectively, whereas in flaviviruses these functions are both associated with the E glycoprotein. Thus, in alphaviruses, the E2 molecules must be removed from the center of the E1 trimer to permit formation of the postfusion complex, requiring the rotation as shown above to externalize E2 from the postfusion trimer. In flaviviruses, the formation of the postfusion trimer is a simpler process and does not necessitate rotation of E prior to fusion.

Transmembrane Region

The improvement in resolution of the 9 Å cryo-EM map compared with the earlier 11 Å map (Zhang et al., 2002b) is particularly useful in the interpretation of the E1 and E2 transmembrane segments. The cryo-EM density clearly shows two helices in each of the four quasi-symmetry-related transmembrane regions. In each of these positions, the E1 helix is bent and does not penetrate beyond the inner phospholipid bilayer, whereas the E2 helix is straight, extends past the bilayer, and interacts with the nucleocapsid core (Figure 6). The possible orientations of the E1 and E2 helices are limited both because of the requirement for small residues in the bend of E1 and because of the proximity of the two helices requiring small residues at the interface between the E1 and E2 glycoproteins. The initial fitting of the E1 and E2 transmembrane helices was in the density surrounding the icosahedral 3-fold axes (monomer 1 in Figure 1). This E1-E2 transmembrane dimer (E1-409 to E1-439 and E2-363 to E2-398) was then fitted independently into the other three quasi-related positions (Table 2). The quasi T = 4 symmetry operators relating the E1-E2 dimers were found to be slightly different compared to those determined for the four E1 ectodomains or the four capsid protein subunits (Table 1), possibly because of the tight packing of the E1-E2 spikes at the 5-fold axes.

The E2 transmembrane helix enters the outer lipid leaflet at residue His363 and emerges past the inner phospholipid leaflet at Cys390. The contiguous stretch of density from the E2 Asn318 carbohydrate moiety (Pletnev et al., 2001) to the N-terminal end of the E2 transmembrane helix could account for the 47 connecting amino acid residues, E2-318 to E2-364. Residues Cys388 and Cys390, which lie in the inner phospholipid head groups, are palmitylated (Strauss and Strauss, 1994). These hydrophobic chains could interact with adjacent aliphatic chains in the lipid bilayer. Residues E2-391 to E2-398 are part of the cytoplasmic region of E2, which

interacts with the nucleocapsid core (see below). Deletions of one or more of these E2 residues affect virus assembly (Hernandez et al., 2000, 2005).

For E1, the transmembrane helix enters the bilayer at residue Trp409 and exits at residue Met433. The carboxyl end of the fitted E1 ectodomain structure (residue Pro383) lies ~25 Å from Trp409, which is an appropriate distance to account for the 24 intervening residues. These residues, which comprise the E1 stem region, form a helix in the postfusion structure (Modis et al., 2004). However, no rod-like density was present in the cryo-EM map in this region. The stretch of residues from E1-415 to E1-420 (Gly-Gly-Ala-Ser-Ser-Leu) was modeled into the E1 density immediately preceding and including the bend in the transmembrane helix. The six carboxy-terminal residues of E1 extend past the inner lipid leaflet into the interior cavity of the virus, but they do not contact the nucleocapsid core. These residues are dispensable for virus assembly ([Barth et al., 1992]; R.J.K., unpublished data).

Chimeric viruses containing RRV E1 and 6K in a SINV background grew poorly compared to wild-type virus, but they produced a second-site revertant at position E2-I380S (Strauss et al., 2002), which is located near the E1-E2 transmembrane interface. Presumably, this isoleucine interferes with dimerization of E1 with E2, and replacement by a serine may facilitate the formation of a favorable interface. Mutation of four glycines to leucines (residues 415, 416, 418, and 423) (Sjoberg and Garoff, 2003) or five glycines to alanines (residues 415, 416, 418, 423, and 435) (Liao and Kielian, 2005) in the E1 transmembrane region of SFV causes no significant change in phenotype, consistent with these residues not being located in the E1-E2 interface. However, when a larger stretch of E1 residues (residues 413–424) was mutated to leucines, there was a decrease in the stability of the heterodimer. Revertants to the leucine mutant were found at residue E1-L417P or E1-L424P, located in the heterodimer interface (Sjoberg and Garoff, 2003).

In contrast to the alphaviruses in which the E1 and E2 glycoproteins each have a single transmembrane helix, the E (envelope) and M (membrane) glycoproteins of flaviviruses each have two transmembrane regions, consistent with sequence analysis predictions and subsequently confirmed by the 9.5 Å structure of dengue virus (Zhang et al., 2003a). The pair of E transmembrane helices forms an antiparallel coiled-coil, whereas the pair of M helices is slightly separated but also antiparallel. There are no interactions between the E and M transmembrane helices in the mature or immature particles. Thus, whereas in alphaviruses the glycoproteins interact with the nucleocapsid core, this does not occur in flaviviruses.

Nucleocapsid Core

The density within the inner lipid leaflet (the nucleocapsid core) is divided into a protein region, a region that probably represents a mixture of protein and RNA, and an internal region probably consisting of only RNA. Although the capsid proteins are related by the same T = 4 quasi-symmetry as the E1 and E2 glycoproteins, their arrangement of pentamers around the 5-fold vertices and hexamers around the icosahedral 2-fold axes are more prominent. The atomic structure of the truncated capsid protein (residues 114–264) was fitted into the cryo-EM density of the protein region of the nucleocapsid core. As with the E1 ectodomain and the transmembrane region, each of the four quasi-equivalent positions was fitted independently. Although the fitted capsid protein has essentially the same T = 4 quasi-symmetry as does the E1 ectodomain, there are slight differences in details (Table 1). Unlike the flaviviruses in which the nucleocapsid core is not in contact with the lipid bilayer or glycoprotein shell, the E2 protein links the surface spikes to the internal nucleocapsid core. As a result, any differences between the four quasi-symmetry-related E1-E2 heterodimers would also influence the capsid protein arrangement. Furthermore, any movement of E2 during receptor binding or fusion might affect the nucleocapsid core.

A difference density map, generated by setting the density of the four quasi-equivalent capsid proteins to zero, showed two regions of extra density. One of these densities is on the outer surface of the nucleocapsid core and extends from the end of the E2 transmembrane helix into the nucleocapsid shell, ending close to the hydrophobic pocket in the capsid protein as predicted (Lee et al., 1996; Skoging et al., 1996) (Figure 7). This places residues E2-400 to E2-402 in the pocket, in agreement with earlier predictions (Lee et al., 1996; Skoging et al., 1996; Wilkinson et al., 2005). No unambiguous density corresponding to residues E2-403 to E2-423 (the carboxy terminus of E2) could be seen exiting the pocket, although it is possible that one region of uninterpreted density might correspond to this E2 polypeptide (see below).

A second region of density in the difference map, ~20 Å long, extends from amino acid 114, the N terminus of the fitted capsid protein, and connects with the protein-RNA region (Zhang et al., 2002b). It is possible to model a 3- or 4-turn helix into this density, roughly accounting for amino acid residues 97–113. The first 100 amino acids of the capsid protein are highly basic and are presumed to bind to the genomic RNA; there is no clear density attributable to these residues. In contrast, flavivirus nucleocapsid cores lack a rigid capsid structure (Zhang et al., 2003a, 2003c), and the flavivirus capsid protein (Dokland et al., 2004; Ma et al., 2004) has basic residues dispersed throughout its ~100 amino acid sequence. Hence, the flavivirus capsid protein is probably dispersed randomly throughout the RNA core, corresponding to the protein-RNA mixed layer in alphaviruse cores.

In the cryo-EM map, there are two regions of uninterpreted low density (0.8σ versus 1.8σ for the regions of difference density described above) that are associated with the inner leaflet of the bilayer and are present at each quasi-symmetry-related position. One of these might be the 6K protein. The low density corresponding to the putative 6K protein may be due to there being less than 240 copies of the 6K molecule per virion (Lusa et al., 1991). The second region of uninterpreted density appears to correspond with residues exiting the hydrophobic pocket on the capsid protein, possibly corresponding to the carboxy end of E2. As this sequence contains some palmitylated residues (Strauss and Strauss, 1994), these aliphatic chains may continue to associate with the lipid membrane. Thus, with this assignment, it has now been possible to produce a fairly complete, pseudoatomic map of an alphavirus.

Experimental Procedures

Virus Cloning, Isolation, and Purification

All site-directed mutations (E2-N318Q, E2-Q46N, and E2-E160N) were made in the TE12 strain of SINV (Lustig et al., 1988) because of its high replication efficiency in BHK cells. To introduce two additional glycosylation sites into the E2 glycoprotein, one amino acid was modified to generate the N-linked glycosylation motif Asn-X-Thr/Ser at E2-Q46N and E2-E160N. Mutant clones were transfected into BHK cells, and the resulting viruses were plaque purified at least twice. The cDNA clones and the genomic RNA were analyzed by DNA sequencing as described previously (Pletnev et al., 2001). The isolation and purification of the mutant viruses were performed as described previously (Pletnev et al., 2001; Strauss et al., 1991).

Electron Microscopy and Image Reconstruction

The single-site deglycosylated virus, E2-N318Q, was used to produce a 9 Å resolution cryo-EM map. Images were recorded on Kodak SO-163 films in a Philips CM200 field emission gun transmission electron microscope (Philips, Eindhoven, The Netherlands) under low-dose conditions (~18 $e^{-}/Å^{2}$) at 38,000 nominal magnification. The actual magnification was calibrated to be 39,220 against an electron density map computed from X-ray diffraction data (Zhang et al., 2003b). Micrographs were digitized on a Zeiss SCAI scanner with 7 μm intervals.

The previously published 11 Å reconstruction map was calculated from 4931 compressed virus images (Zhang et al., 2002b) (pixel size was 3.57 Å) of the same data set. For the 9 Å image reconstruction, a total of 10,868 particles were selected from 27 micrographs recorded at defocus levels ranging between 1.10 and 2.58 μm underfocus. The Fourier transform of each image was modified to correct the microscope contrast transfer function. Uncompressed images, whose pixel size was 1.78 Å, were used for the final refinement of orientation and origin parameters (Ji et al., 2003). This new refinement algorithm, in conjunction with parallelized computer programs, improved the accuracy and speed of the computations, allowing more images to be included in the reconstruction process. The final map was computed from 7085 images. Phase agreement (<50°) and Fourier shell correlation coefficients (FCS) (>0.5) for all protein-occupying regions (nucleocapsid core, lipid bilayer, and outer glycoprotein shell) indicated that the resolution of the final map was ~9.0 Å. If each region were to be assessed individually, the lipid membrane region (which contains the least amount of protein) would have the worst FCS values at most spatial frequencies, whereas the outer glycoprotein shell would have the best FCS values.

Sample preparation, data collection, and image processing of the hyperglycosylated mutants were carried out essentially as described previously (Baker et al., 1999; Mukhopadhyay et al., 2002; Pletnev et al., 2001; Zhang et al., 2002b). The effective resolutions of the cryo-EM reconstructions obtained from these mutant virus samples were 18 Å for E2-Q46N and E2-E216N and 26 Å for E2-E160N. Difference maps were generated as described previously (Pletnev et al., 2001; Zhang et al., 2002a) by subtracting the density map of native SINV from each of these maps to locate the carbohydrate moieties associated with E2-46, E2-160, and E2-216.

Interpretation of Cryo-EM Reconstructions

SINV E1 was modeled from the partially refined coordinates of the homologous SFV E1 structure of residues 1–383 (generously provided by F. Rey). The transmembrane regions of E1 (residues 408– 439) and E2 (residues 363–398) were modeled as α helices.

The E1 glycoprotein ectodomain was fitted as two, independent, rigid bodies, domains I + II and domain III. All nonhydrogen atoms of the atomic structures were used in the fitting process. The high quality of the cryo-EM map at 9 Å resolution allowed each of the T = 4 quasi-symmetry-related monomers to be fitted independently. The quality of each fit was monitored by the *sumf*, clash, and negative density parameters used by the EMfit program (Rossmann, 2000; Rossmann et al., 2001).

Acknowledgments

We thank Rob Ashmore and Chuan Xiao for developing and maintaining numerous computer and interactive graphic programs, and Ying Zhang and Petr Leiman for helpful discussions and Jolanda Smit for purified SFV. The work was supported in part by National Institutes of Health (NIH) Program Project Grant (AI55672) to T.S.B., R.J.K., M.G.R., and J.H.S., and by NIH grants to T.S.B. (GM R37-033050), R.J.K. (GM56279), and J.H.S. (AI20612).

References

Baker TS, Olson NH, Fuller SD. Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. Microbiol Mol Biol Rev 1999;63:862–922. [PubMed: 10585969]

Barth BU, Suomalainen M, Liljestrom P, Garoff H. Alphavirus assembly and entry: role of the cytoplasmic tail of the E1 spike subunit. J Virol 1992;66:7560–7564. [PubMed: 1331539]

Bressanelli S, Stiasny K, Allison SL, Stura EA, Duquerroy S, Lescar J, Heinz FX, Rey FA. Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. EMBO J 2004;23:728–738. [PubMed: 14963486]

Caspar DLD, Klug A. Physical principles in the construction of regular viruses. Cold Spring Harb Symp Quant Biol 1962;27:1–24. [PubMed: 14019094]

Cheng RH, Kuhn RJ, Olson NH, Rossmann MG, Choi HK, Smith TJ, Baker TS. Nucleocapsid and glycoprotein organization in an enveloped virus. Cell 1995;80:621–630. [PubMed: 7867069]

Choi HK, Tong L, Minor W, Dumas P, Boege U, Rossmann MG, Wengler G. Structure of Sindbis virus core protein reveals a chymotrypsin-like serine proteinase and the organization of the virion. Nature 1991;354:37–43. [PubMed: 1944569]

Crill WD, Roehrig JT. Monoclonal antibodies that bind to domain III of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to Vero cells. J Virol 2001;75:7769–7773. [PubMed: 11462053]

Davis NL, Pence DF, Meyer WJ, Schmaljohn AL, Johnston RE. Alternative forms of a strain-specific neutralizing antigenic site on the Sindbis virus E2 glycoprotein. Virology 1987;161:101–108. [PubMed: 2445100]

Dokland T, Walsh M, Mackenzie JM, Khromykh AA, Ee KH, Wang S. West nile virus core protein; tetramer structure and ribbon formation. Structure 2004;12:1157–1163. [PubMed: 15242592]

Ellenberger TE, Brandl CJ, Struhl K, Harrison SC. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystal structure of the protein-DNA complex. Cell 1992;71:1223–1237. [PubMed: 1473154]

Gaedigk-Nitschko K, Schlesinger MJ. The Sindbis virus 6K protein can be detected in virions and is acylated with fatty acids. Virology 1990;175:274–281. [PubMed: 2408229]

Gibbons DL, Erk I, Reilly B, Navaza J, Kielian M, Rey FA, Lepault J. Visualization of the target-membrane-inserted fusion protein of Semliki Forest virus by combined electron microscopy and crystallography. Cell 2003;114:573–583. [PubMed: 13678581]

Gibbons DL, Vaney MC, Roussel A, Vigouroux A, Reilly B, Lepault J, Kielian M, Rey FA. Conformational change and protein-protein interactions of the fusion protein of Semliki Forest virus. Nature 2004;427:320–325. [PubMed: 14737160]

Heil ML, Albee A, Strauss JH, Kuhn RJ. An amino acid substitution in the coding region of the E2 glycoprotein adapts Ross River virus to utilize heparan sulfate as an attachment moiety. J Virol 2001;75:6303–6309. [PubMed: 11413296]

Hernandez R, Lee H, Nelson C, Brown DT. A single deletion in the membrane-proximal region of the Sindbis virus glycoprotein E2 endodomain blocks virus assembly. J Virol 2000;74:4220–4228. [PubMed: 10756035]

Hernandez R, Ferreira D, Sinodis C, Litton K, Brown DT. Single amino acid insertions at the junction of the Sindbis virus E2 transmembrane domain and endodomain disrupt virus envelopment and alter infectivity. J Virol 2005;79:7682–7697. [PubMed: 15919921]

Ji, Y.; Marinescu, DC.; Zhang, W.; Baker, TS. Orientation refinement of virus structures with unknown symmetry. Paper presented at: IEEE Computer Society; Los Alamitos, California. 2003.

Kuhn RJ, Zhang W, Rossmann MG, Pletnev SV, Corver J, Lenches E, Jones CT, Mukhopadhyay S, Chipman PR, Strauss EG, et al. Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. Cell 2002;108:717–725. [PubMed: 11893341]

Lee S, Owen KE, Choi HK, Lee H, Lu G, Wengler G, Brown DT, Rossmann MG, Kuhn RJ. Identification of a protein binding site on the surface of the alphavirus nucleocapsid and its implication in virus assembly. Structure 1996;4:531–541. [PubMed: 8736552]

Lescar J, Roussel A, Wien MW, Navaza J, Fuller SD, Wengler G, Rey FA. The fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. Cell 2001;105:137–148. [PubMed: 11301009]

Liao M, Kielian M. The conserved glycine residues in the transmembrane domain of the Semliki Forest virus fusion protein are not required for assembly and fusion. Virology 2005;332:430–437. [PubMed: 15661173]

Lusa S, Garoff H, Liljestrom P. Fate of the 6K membrane protein of Semliki Forest virus during virus assembly. Virology 1991;185:843–846. [PubMed: 1962454]

Lustig S, Jackson AC, Hahn CS, Griffin DE, Strauss EG, Strauss JH. Molecular basis of Sindbis virus neurovirulence in mice. J Virol 1988;62:2329–2336. [PubMed: 2836615]

Ma L, Jones CT, Groesch TD, Kuhn RJ, Post CB. Solution structure of dengue virus capsid protein reveals another fold. Proc Natl Acad Sci USA 2004;101:3414–3419. [PubMed: 14993605]

Mancini EJ, Clarke M, Gowen BE, Rutten T, Fuller SD. Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. Mol Cell 2000;5:255–266. [PubMed: 10882067]

Mandl CW, Allison SL, Holzmann H, Meixner T, Heinz FX. Attenuation of tick-borne encephalitis virus by structure-based site-specific mutagenesis of a putative flavivirus receptor binding site. J Virol 2000;74:9601–9609. [PubMed: 11000232]

Meyer WJ, Johnston RE. Structural rearrangement of infecting Sindbis virions at the cell surface: mapping of newly accessible epitopes. J Virol 1993;67:5117–5125. [PubMed: 7688818]

Modis Y, Ogata S, Clements D, Harrison SC. Structure of the dengue virus envelope protein after membrane fusion. Nature 2004;427:313–319. [PubMed: 14737159]

Mukhopadhyay S, Chipman PR, Hong EM, Kuhn RJ, Rossmann MG. In vitro-assembled alphavirus core-like particles maintain a structure similar to that of nucleocapsid cores in mature virus. J Virol 2002;76:11128–11132. [PubMed: 12368355]

Paredes AM, Brown DT, Rothnagel R, Chiu W, Schoepp RJ, Johnston RE, Prasad BV. Three-dimensional structure of a membrane-containing virus. Proc Natl Acad Sci USA 1993;90:9095–9099. [PubMed: 8415660]

Paredes AM, Heidner H, Thuman-Commike P, Prasad BV, Johnston RE, Chiu W. Structural localization of the E3 glycoprotein in attenuated Sindbis virus mutants. J Virol 1998;72:1534–1541. [PubMed: 9445057]

Paredes A, Alwell-Warda K, Weaver SC, Chiu W, Watowich SJ. Venezuelan equine encephalomyelitis virus structure and its divergence from old world alphaviruses. J Virol 2001;75:9532–9537. [PubMed: 11533216]

Pletnev SV, Zhang W, Mukhopadhyay S, Fisher BR, Hernandez R, Brown DT, Baker TS, Rossmann MG, Kuhn RJ. Locations of carbohydrate sites on alphavirus glycoproteins show that E1 forms an icosahedral scaffold. Cell 2001;105:127–136. [PubMed: 11301008]

Rey FA, Heinz FX, Mandl C, Kunz C, Harrison SC. The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. Nature 1995;375:291–298. [PubMed: 7753193]

Rossmann MG. Fitting atomic models into electron-microscopy maps. Acta Crystallogr D Biol Crystallogr 2000;56:1341–1349. [PubMed: 10998631]

Rossmann MG, Bernal R, Pletnev SV. Combining electron microscopic with x-ray crystallographic structures. J Struct Biol 2001;136:190–200. [PubMed: 12051899]

Sjoberg M, Garoff H. Interactions between the transmembrane segments of the alphavirus E1 and E2 proteins play a role in virus budding and fusion. J Virol 2003;77:3441–3450. [PubMed: 12610119]

Skoging U, Vihinen M, Nilsson L, Liljestrom P. Aromatic interactions define the binding of the alphavirus spike to its nucleocapsid. Structure 1996;4:519–529. [PubMed: 8736551]

Smith TJ, Cheng RH, Olson NH, Peterson P, Chase E, Kuhn RJ, Baker TS. Putative receptor binding sites on alphaviruses as visualized by cryoelectron microscopy. Proc Natl Acad Sci USA 1995;92:10648–10652. [PubMed: 7479858]

Strauss JH, Strauss EG. The alphaviruses: gene expression, replication, and evolution. Microbiol Rev 1994;58:491–562. [PubMed: 7968923]

Strauss EG, Stec DS, Schmaljohn AL, Strauss JH. Identification of antigenically important domains in the glycoproteins of Sindbis virus by analysis of antibody escape variants. J Virol 1991;65:4654–4664. [PubMed: 1714515]

Strauss EG, Lenches EM, Strauss JH. Molecular genetic evidence that the hydrophobic anchors of glycoproteins E2 and E1 interact during assembly of alphaviruses. J Virol 2002;76:10188–10194. [PubMed: 12239293]

von Bonsdorff CH, Harrison SC. Sindbis virus glycoproteins form a regular icosahedral surface lattice. J Virol 1975;16:141–145. [PubMed: 48559]

von Bonsdorff CH, Harrison SC. Hexagonal glycoprotein arrays from Sindbis virus membranes. J Virol 1978;28:578–583. [PubMed: 722862]

Wilkinson TA, Tellinghuisen TL, Kuhn RJ, Post CB. Association of Sindbis virus capsid protein with phospholipid membranes and the E2 glycoprotein: implications for alphavirus assembly. Biochemistry 2005;44:2800–2810. [PubMed: 15723524]

Zhang W, Fisher BR, Olson NH, Strauss JH, Kuhn RJ, Baker TS. Aura virus structure suggests that the T = 4 organization is a fundamental property of viral structural proteins. J Virol 2002a;76:7239–7246. [PubMed: 12072523]

Zhang W, Mukhopadhyay S, Pletnev SV, Baker TS, Kuhn RJ, Rossmann MG. Placement of the structural proteins in Sindbis virus. J Virol 2002b;76:11645–11658. [PubMed: 12388725]

Zhang W, Chipman PR, Corver J, Johnson PR, Zhang Y, Mukhopadhyay S, Baker TS, Strauss JH, Rossmann MG, Kuhn RJ. Visualization of membrane protein domains by cryo-electron microscopy of dengue virus. Nat Struct Biol 2003a;10:907–912. [PubMed: 14528291]

Zhang X, Walker SB, Chipman PR, Nibert ML, Baker TS. Reovirus polymerase lambda 3 localized by cryo-electron microscopy of virions at a resolution of 7.6 Å. Nat Struct Biol 2003b;10:1011–1018. [PubMed: 14608373]

Zhang Y, Corver J, Chipman PR, Zhang W, Pletnev SV, Sedlak D, Baker TS, Strauss JH, Kuhn RJ, Rossmann MG. Structures of immature flavivirus particles. EMBO J 2003c;22:2604–2613. [PubMed: 12773377]

Zhang Y, Zhang W, Ogata S, Clements D, Strauss JH, Baker TS, Kuhn RJ, Rossmann MG. Conformational changes of the flavivirus E glycoprotein. Structure 2004;12:1607–1618. [PubMed: 15341726]

Zhang W, Heil M, Kuhn RJ, Baker TS. Heparin binding sites on Ross River virus revealed by electron cryo-microscopy. Virology 2005;332:511–518. [PubMed: 15680416]

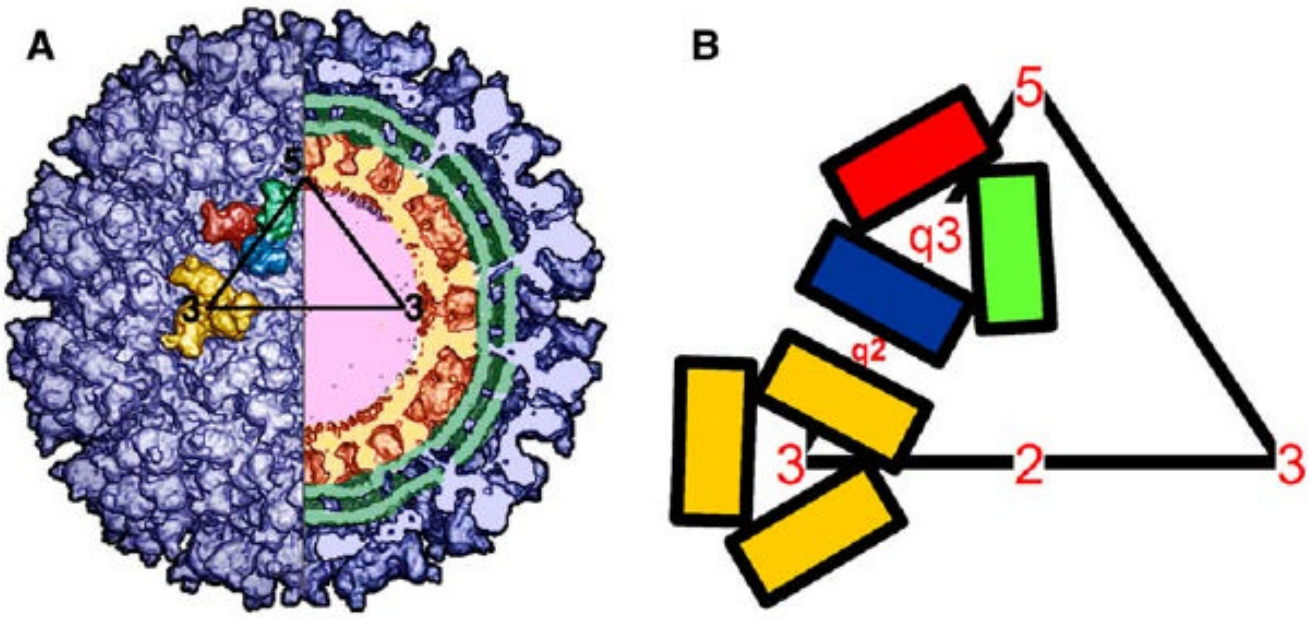

Figure 1. The T = 4 Quasi-Symmetry of SINV

(A) Surface-shaded view of Sindbis virus shown at 9 Å resolution. Left half: The four independent E2 monomers related by T = 4 quasi-symmetry are colored yellow (i3), blue (m4), green (m5), and red (m6). The three yellow monomers are identical in structure and constitute the spikes at the icosahedral 3-fold axes. The blue, green, and red monomers are similar, but not identical, in structure and constitute the spikes at the quasi-3-fold axes. The triangle delineates 1 of the 60 icosahedral asymmetric units. Right half: A cross-section of the virus particle showing the overall organization of the particle. The nucleocapsid is colored red, the lipid bilayer is green, and the glycoproteins are shown in blue.

(B) Diagram of the T = 4 quasi-symmetry-related monomers of the E1 protein. The yellow (i3), blue (m4), green (m5), and red (m6), monomers of E1 form heterodimers with the corresponding colored E2 molecules.

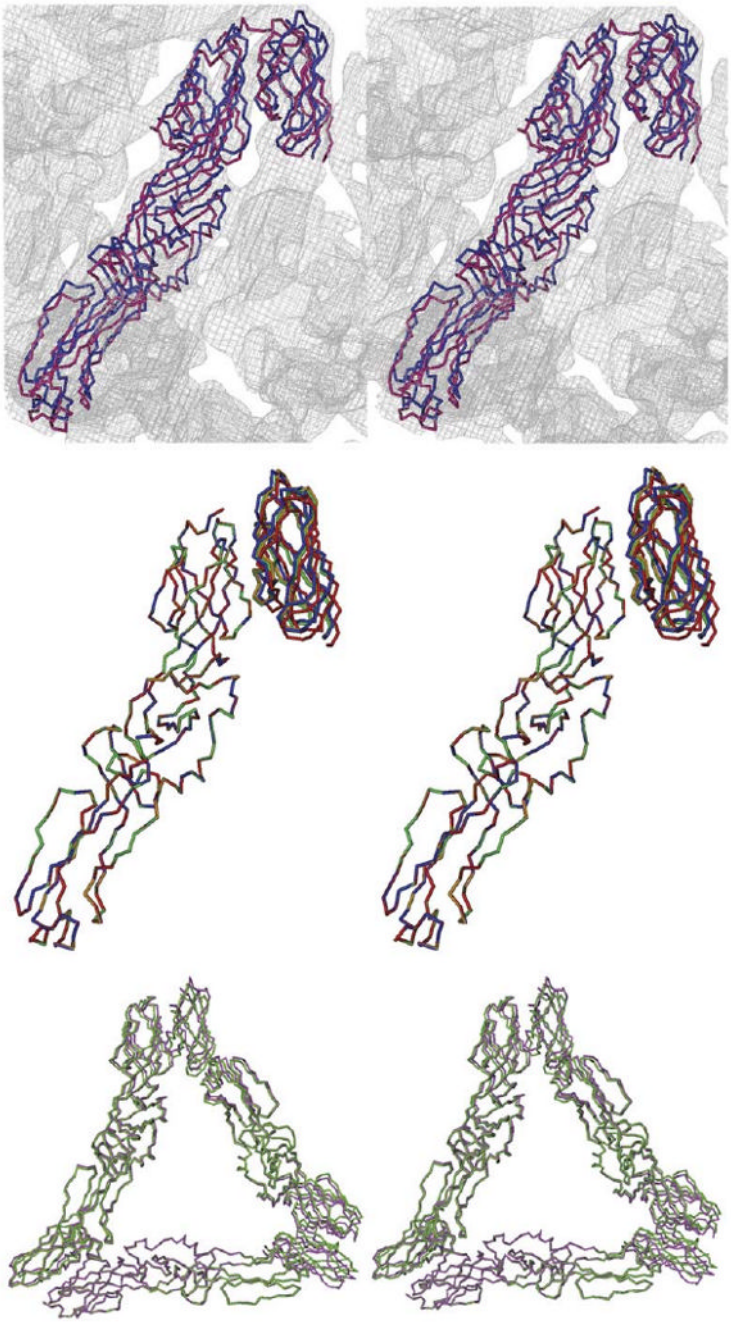

Figure 2. Stereodiagrams Showing the Fitted E1 Glycoprotein

Top: The Cα backbone of the E1 monomer fitted into the 9 Å resolution cryo-EM map is shown at the spike around the icosahedral 3-fold axis. E1 is shown in red when fitted as an entire molecule, and it is shown in blue when fitted as two independent rigid bodies (domains I+II and domain III). Center: Domains I and II of the four quasi-symmetry-related E1 monomers shown in yellow, blue, green, and red (as in Figure 1) are superimposed, showing the relative position of domain III. The red monomer, which is closest to the 5-fold axis (see Figure 1), has the greatest deviation from the other monomers. Bottom: The E1 trimer surrounding the icosahedral 3-fold axes (green) superimposed on the E1 trimer surrounding the quasi-3-fold axes (purple) is shown.

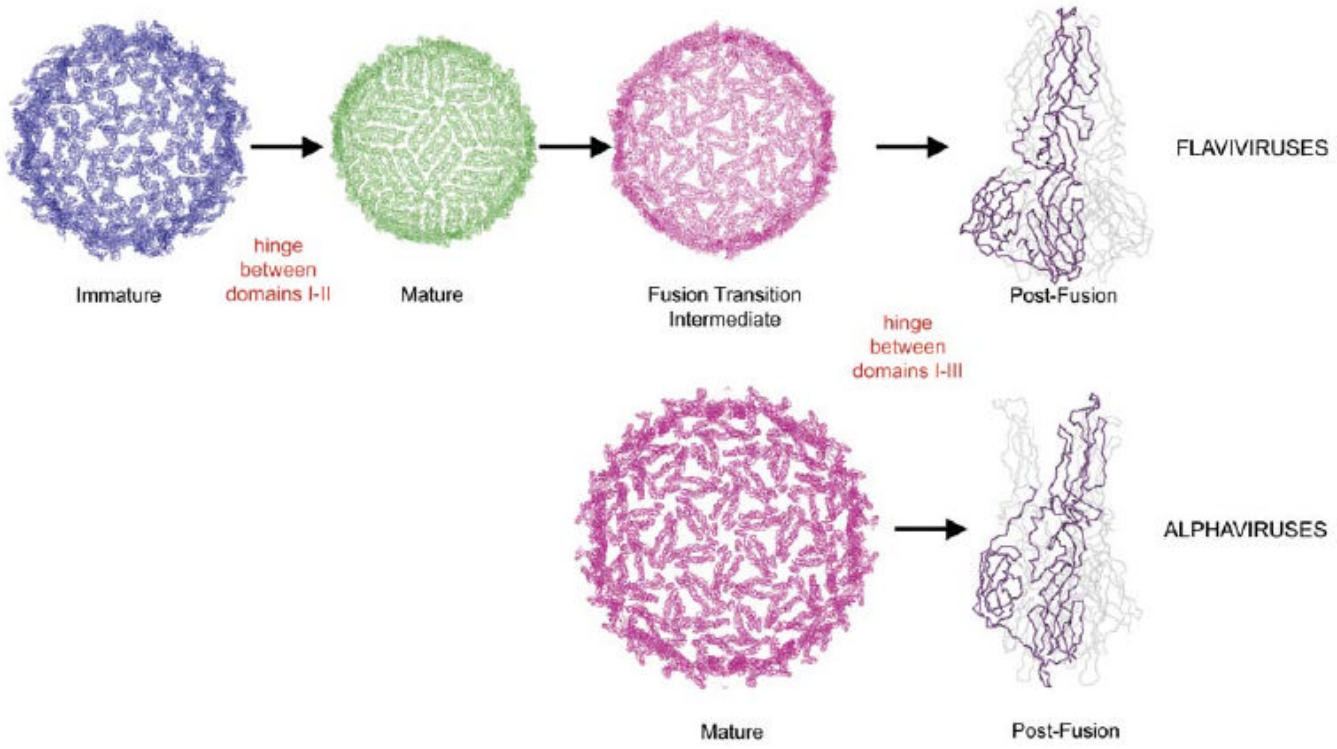

Figure 3. Steps in the Maturation and Fusion Processes Involving the Flavivirus E and Alphavirus E1 Proteins

Flavivirus E is shown at the top; alphavirus E1 is shown at the bottem. The maturation process of flaviviruses requires extensive conformational changes of the E protein as the spikes of the immature particle rearrange to become a smooth-surfaced mature particle. Flaviviruses have been postulated to have an intermediate state before fusion and formation of E trimers. Mature alphaviruses have an arrangement of their E1 protein similar to the proposed fusion transition intermediate of flaviviruses. The E1 trimeric postfusion spike is similar to that of flaviviruses.

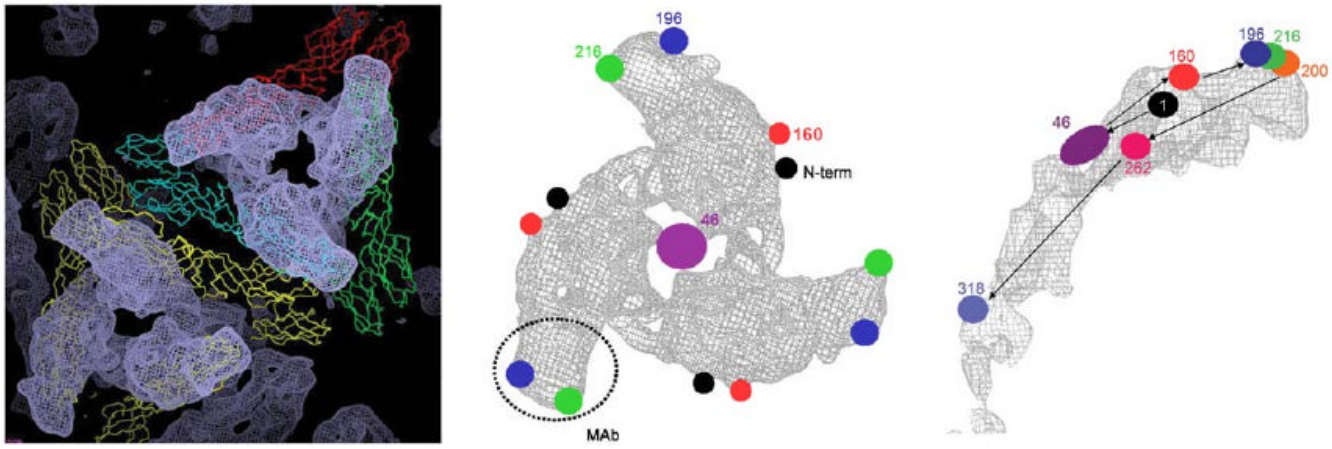

Figure 4. Mapping the E2 Ectodomain Markers

Left: The E2 difference density for two surface spikes is shown in gray. The Cα backbones of the four quasi-symmetry-related E1 monomers are shown in yellow, blue, green, and red. Center: A top view of the E2 density of one spike looking down the 3-fold axis (shown as the larger purple spot corresponding to the three merged carbohydrate moieties at position 46). Right: A side view of one E2 molecule. Center and Right: The markers on the E2 glycoprotein that correspond to carbohydrate moieties at positions 46, 160, 196, 200, 216, 262, and 318 are shown in purple, red, blue, orange, green, pink, and light blue, respectively. Position 216 (green) was also identified with a cryo-EM map of a Fab-Ross River virus complex. Position 1 (black) was determined by a difference map between SFV (which contains E3) and SINV (which lacks E3).

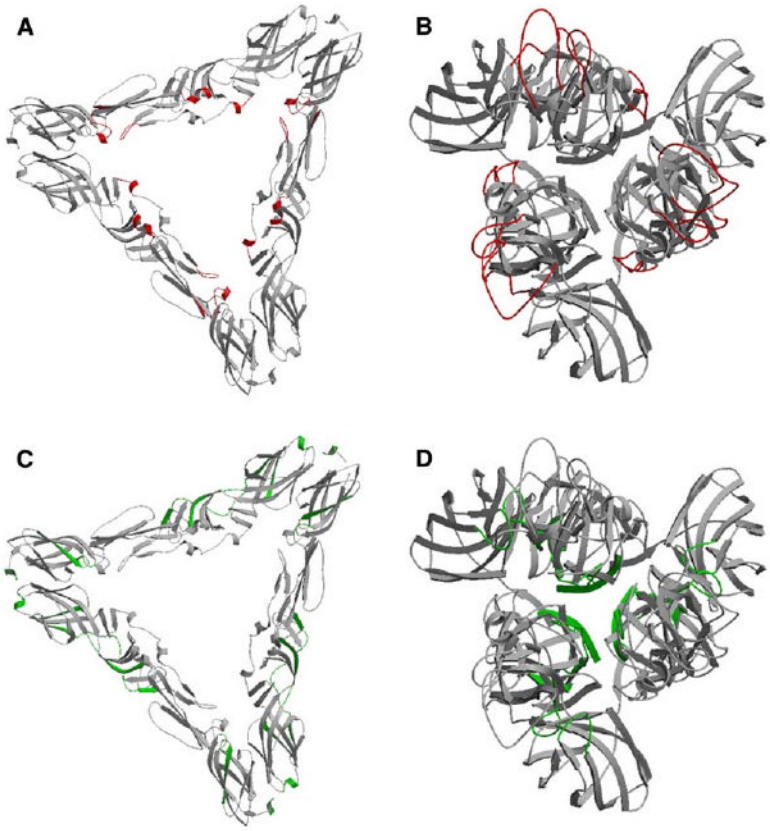

Figure 5. Identification of Putative E1 Residues that Contact E2 in the Pre- and Postfusion States

(A) The icosahedral 3-fold trimer of E1 molecules (gray) is shown, with residues that contact the E2 density in red.

(B) The location of these same residues mapped onto the E1 trimer in the postfusion complex is shown. A majority of these residues are on the external surface of the postfusion trimer.

(C) The icosahedral 3-fold trimer of E1 molecules (gray) is shown with residues (green) that make intermonomer contacts in the postfusion state.

(D) The postfusion trimer shown with residues (green) that make intermonomer contacts. These observations show that the E1 molecules are rotated about their long axes during the transition from the prefusion to the postfusion state, thereby presumably moving the E2 molecules from the inside to the outside of the E1 trimer while the E1-E2 heterodimer dissociates.

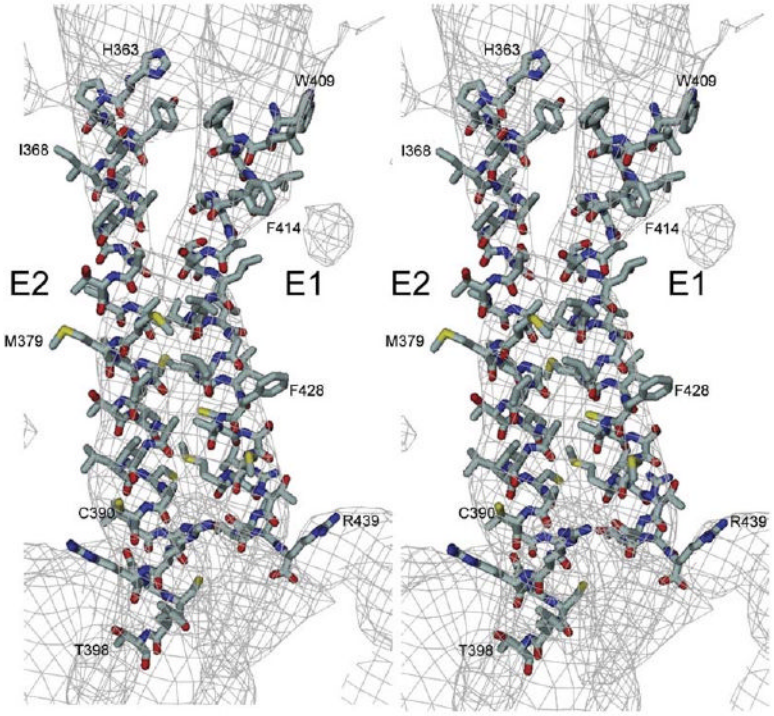

Figure 6. The E1 and E2 Transmembrane Helices

Stereodiagram of E1 residues from 409 to 439 and E2 residues from 363 to 398 fitted into their transmembrane densities. The bend in the E1 density around residue Ala417 is present in all four quasi-symmetry-related positions.

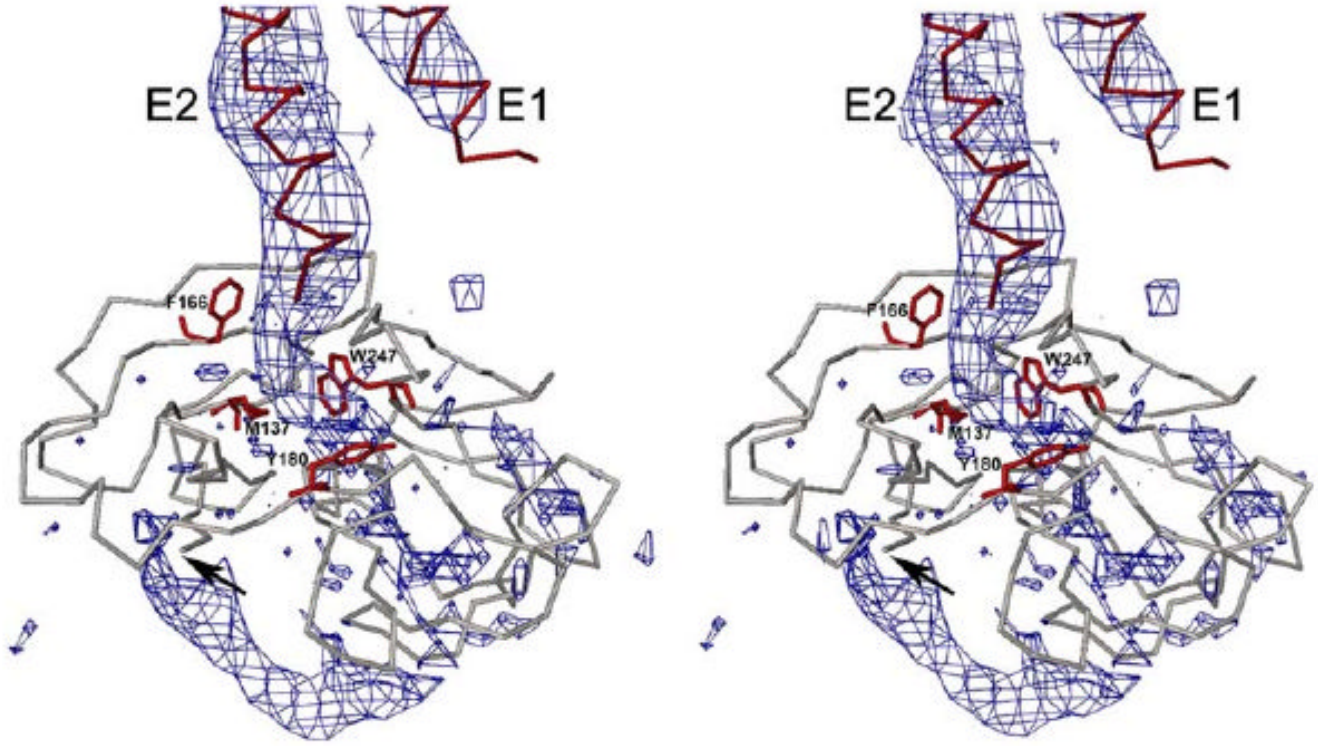

Figure 7. Stereodiagram Showing the Remaining Density after Subtracting the Density Corresponding to the Fitted X-Ray Crystallographic Capsid Protein for Residues 114–265

The residues in the hydrophobic pocket that constitutes the E2 capsid protein binding site are shown in red. An arrow identifies residue 114 at the N terminus of the known capsid structure. The difference density that ends at the arrow probably corresponds to residues 97–113 of the capsid protein.

Table 1

Fitting of the Crystallographic Structures into the Cryo-EM Map by Using the Program EMfit

Molecule[a]	Quasi-Symmetry[b]	Position[c]	*sumf* Total	Clash	Negative Density	*sumf* in Individual Domains[d]		
						Domain I	Domain II	Domain III
E1 (entire)	A	All	41.5	0.2	4.4			
		i3				35.8	42.9	43.3
		m4				36.8	43.5	43.1
		m5				39.5	45.1	42.3
		m6				37.1	44.7	40.2
E1 (DI+DII plus DIII)	B	All	43.8	0.1	2.7			
E1 (DI+DII)	B	i3	43.2		3.1	37.6	46.2	
		m4	42.2		3.4	37.4	44.8	
		m5	44.3		2.7	39.7	46.8	
		m6	43.1		3.1	37.4	46.2	
E1 (DIII)	B	i3	46.5		1.2			46.5
		m4	47.8		0.3			47.8
		m5	45.9		0.0			45.9
		m6	44.1		1.5			44.1
GCN4 helices	A	All	30.8	0.0	13.8			
		i3				32.2	34.8	
		m4				27.5	33.0	
		m5				22.7	23.7	
		m6				30.8	32.2	
E1 and E2 TM	B	All	32.9	0.0	11.5			
		i3				34.2	36.7	
		m4				32.7	33.6	
		m5				30.3	31.5	
		m6				27.7	33.0	
SCP	A	All	41.2	0.3	5.1			

Molecule[a]	Quasi-Symmetry[b]	Position[c]	*sumf* Total	Clash	Negative Density	*sumf* in Individual Domains[d]		
						Domain I	Domain II	Domain III
		i3				42.3		
		m4				40.2		
		m5				41.8		
		m6				40.5		
SCP	B	All	43.3	0.2	3.0			
		i3	44.9		2.8	44.9		
		m4	44.0		2.3	44.0		
		m5	43.5		2.8	43.5		
		m6	40.6		3.3	40.6		

[a] All nonhydrogen atoms were used, except for GCN4 (Ellenberger et al., 1992), in which only Cα atoms of residues 250–277, PDB 1YSA, were used. TM, transmembrane; SCP, Sindbis capsid protein.

[b] Quasi-symmetry operators used from Zhang et al. (2002b) (see Table 2 for values) are indicated by “A.” Quasi-symmetry operators determined by individually fitting each of the quasi-equivalent monomers and then determining the symmetry operators (Table 2) are indicated by “B”.

[c] “All” indicates that more than 28 molecules were used while fitting in order to determine the total *sumf* (%), the number of clashes (%), and the negative density (%) as defined in Rossmann et al. (2001). The individual monomers i3, m4, m5, and m6 correspond to the positions in Figure 1.

[d] For the E1 and E2 TM region, domains I and II correspond to E1 and E2, respectively.

Table 2

Polar Angle Defining Quasi-Symmetric Monomers

	Quasi-2-Fold			Quasi-3-Fold								
	Rotation of Molecule 1 to 4			Rotation of Molecule 4 to 5			Rotation of Molecule 5 to 6			Rotation of Molecule 6 to 4		
Molecule	**Ψ**	**Φ**	**K**	**Ψ**	**Φ**	**K**	**Ψ**	**Φ**	**K**	**Ψ**	**Φ**	**K**
All proteins[a]	73.9	−81.2	180.0	79.6	−72.6	120.0	79.6	−72.6	120.0	79.6	−72.6	120.0
E1 (DI+DII plus DIII)	74.5	−83.4	181.4	81.8	−71.3	124.3	78.4	−72.6	118.9	79.1	−68.6	116.9
E1 (DI+DII)	73.4	−79.4	181.4	80.9	−71.9	122.8	80.3	−72.8	118.2	79.8	−71.8	119.0
E1 (DIII)	76.0	−84.3	182.9	80.7	−76.2	122.5	74.0	−74.8	118.3	78.7	−69.6	120.0
E1 and E2 (TM)[b]	74.3	−81.9	179.1	80.5	−75.4	123.3	79.0	−77.1	113.4	78.5	−74.9	123.5
SCP[c]	73.9	−80.0	178.6	79.7	−71.3	119.2	78.2	−71.3	117.5	78.9	−70.0	123.3

Polar angles are defined as in Rossmann (2000), with respect to the Cartesian axes shown in Figure 1.

[a] Quasi-symmetry used in Zhang et al. (2002b) for all proteins fitted.

[b] TM, transmembrane.

[c] SCP, Sindbis capsid protein.